

OPEN

Characterisation of HIV-1 Molecular Epidemiology in Nigeria: Origin, Diversity, Demography and Geographic Spread

Jamirah Nazziwa¹, Nuno Rodrigues Faria², Beth Chaplin³, Holly Rawizza³, Phyllis Kanki³, Patrick Dakum^{4,5}, Alash'le Abimiku^{4,5}, Man Charurat⁵, Nicaise Ndembu^{4,5} & Joakim Esbjörnsson^{1,6*}

Nigeria has the highest number of AIDS-related deaths in the world. In this study, we characterised the HIV-1 molecular epidemiology by analysing 1442 HIV-1 *pol* sequences collected 1999–2014 from four geopolitical zones in Nigeria using state-of-the-art maximum-likelihood and Bayesian phylogenetic analyses. The main circulating forms were the circulating recombinant form (CRF) 02_AG (44% of the analysed sequences), CRF43_02G (16%), and subtype G (8%). Twenty-three percent of the sequences represented unique recombinant forms (URFs), whereof 37 (11%) could be grouped into seven potentially novel CRFs. Bayesian phylodynamic analysis suggested that five major Nigerian HIV-1 sub-epidemics were introduced in the 1960s and 1970s, close to the Nigerian Civil War. The analysis also indicated that the number of effective infections decreased in Nigeria after the introduction of free antiretroviral treatment in 2006. Finally, Bayesian phylogeographic analysis suggested gravity-like dynamics in which virus lineages first emerge and expand within large urban centers such as Abuja and Lagos, before migrating towards smaller rural areas. This study provides novel insight into the Nigerian HIV-1 epidemic and may have implications for future HIV-1 prevention strategies in Nigeria and other severely affected countries.

Thirty-eight years after the first AIDS case was described, HIV-1 remains a major public health problem that affects approximately 36.7 (30.8–42.9) million people globally^{1,2}. Sub-Saharan Africa accounts for approximately 70% of all infections. In addition, HIV-1 represent one of the most fast-evolving pathogens known to science. This has led to the classification of HIV-1 into subtypes, sub-subtypes and circulating recombinant forms (CRFs). Of relevance, previous reports have suggested that the genetic composition of the infecting HIV-1 strain may influence transmission, disease progression, virus-host interactions, antiretroviral treatment responses, and vaccine development^{3–12}. Moreover, it has been suggested that targeted HIV-1 prevention may be a cost-effective way to decrease the number of new HIV-1 infections – not least in low-and-middle-income countries¹³. Planning of such HIV-1 intervention programs will require detailed information about both the characteristics and sources of new infections.

Nigeria is the most populous country in Sub-Saharan Africa, has the second highest number of HIV-1 infected persons in the world, and the highest number of annual AIDS-related deaths^{14,15}. HIV-1 serological surveys in Nigeria were initiated in 1991, at which time the adult prevalence was 1.8% (760000 people)¹⁶. This figure increased to 5.8% by 2001 (2.6 million), before declining to 2.8% in 2017 (3.1 million)^{2,16}. Previous reports have identified subtype G and the CRF02_AG as the most common HIV-1 variants in Nigeria^{17–20}. In addition, the CRF43_02G was recently shown to be highly prevalent in the capital Abuja²¹. However, the estimates of each strain's contribution to the Nigerian HIV-1 epidemic have varied considerably, with frequency estimates ranging from 22–50% for subtype G, and 19–60% for the CRF02_AG^{18–25}. These variations could have several reasons,

¹Department of Translational Medicine, Lund University, Lund, Sweden. ²Department of Zoology, University of Oxford, Oxford, United Kingdom. ³Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, USA. ⁴Institute of Human Virology Nigeria, Abuja, Nigeria. ⁵Institute of Human Virology, University of Maryland School of Medicine, Baltimore, USA. ⁶Nuffield Department Medicine, University of Oxford, Oxford, United Kingdom. *email: joakim.esbjornsson@med.lu.se

Subtype/ CRF ^a	N ^c	%	Geography ^d																		
			Southwest							North Central					North East & North West						
			LAG	IBA	OND	OSO	OYO	ENU	Total	ABV	JOS	MIN	NC	Total	KAD	DAM	MAI	KAN	YOL	Total	
A1	24	1.7	9	2	0	0	0	0	0	11	7	5	0	0	12	0	0	1	0	0	1
B	4	0.3	1	1	0	0	0	0	0	2	2	0	0	0	2	0	0	0	0	0	0
C	14	1.0	2	1	0	0	2	0	0	5	6	2	0	1	9	0	0	0	0	0	0
CRF02_AG	636	44.1	176	27	0	1	18	0	0	222	284	94	0	10	388	13	0	11	2	0	26
CRF06_cpx	64	4.4	18	2	0	0	0	0	0	20	36	6	0	1	43	0	0	1	0	0	1
CRF09_cpx	1	0.1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
CRF11_cpx	4	0.3	1	0	0	0	0	0	0	1	2	1	0	0	3	0	0	0	0	0	0
CRF18_cpx	1	0.1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
CRF19_cpx	1	0.1	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0
CRF43_02G	236	16.4	33	11	1	0	0	0	0	45	138	34	0	0	172	8	0	8	3	0	19
CRF49_cpx	1	0.1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
D	9	0.6	2	0	0	0	0	0	0	2	4	2	0	0	6	0	0	1	0	0	1
G	119	8.3	24	6	0	0	0	0	0	30	55	25	1	0	81	4	1	3	0	0	8
URF ^b	328	22.7	78	13	0	0	2	1	1	94	163	46	0	3	212	9	0	11	0	2	22
Total	1442	100	346	64	1	1	22	1	1	435	697	216	1	15	929	34	1	35	5	2	78

Table 1. Proportion of subtype/CRF in the Nigerian dataset of previously published and new sequences collected in the period 1999–2013. ^aCRF, Circulating recombinant forms; ^bURF, Unique recombinant forms; ^cN, number of sequences; ^dAbbreviations for the different cities: LAG, Lagos; IBA, Ibadan; OND, Ondo; OSO, Osogbo; OYO, Oyo; ENU, Enugu; ABV, Abuja; JOS, Jos; MIN, Minna; NC, Other North central cities; KAD, Kaduna; DAM, Damaturu; MAI, Maiduguri; KAN, Kano; YOL, Yola.

such as differences between geographic areas, transmission groups, or employed subtyping tools^{26,27}. Hence, a clearer picture of the HIV-1 subtype/CRF distributions and spread in Nigeria is needed. The objective of the current study was therefore to perform the first nationwide analysis of the molecular epidemiology of HIV-1 in Nigeria.

Methods

Sequence dataset. We analyzed 366 previously unpublished HIV-1 *pol* sequences (positions in HXB2 K03455: 2253–3364) collected in Abuja, Nigeria between 2006–2011, together with all Nigerian *pol* sequences from the corresponding genetic region available in the Los Alamos HIV-1 sequence database (N = 1076, <http://www.hiv.lanl.gov/>, Table 1). All sequences without date and location of sampling were excluded. However, contact with all relevant research centers was made in order to obtain such missing information.

Subtype determination. The Nigerian *pol* sequences were aligned with all M group lineages (A–K + Recombinants) of the LANL 2010 Reference Sequence Dataset (<http://www.hiv.lanl.gov/>) using the Clustal algorithm, followed by manual editing in MEGA6^{28,29}. The HIV-1 subtype/CRF classification was determined by manual assignment with maximum-likelihood (ML) phylogenetic analysis in Garli v0.98³⁰ using the General Time Reversible (GTR) substitution model. Branch support was determined in PhyML 3.0 using the approximate likelihood ratio test with the Shimodaira–Hasegawa-like procedure (aLRT-SH)³¹. Branches with aLRT-SH support >0.90 were considered statistically supported³².

Recombination analysis. Putative unique recombinant forms (URFs) and sequences that were difficult to subtype were analysed by Bootscan in Simplot³³. Briefly, *pol* sequences were aligned with the LANL 2010 Reference sequences for subtypes for G, CRF4302G and CRF02AG as putative parental sequences. Recombination breakpoints were identified using a sliding window size of 300 bp and step size of 50 bp. URFs consisting of non-subtype G related forms were excluded from the Simplot analysis.

In order to define the structure and distribution of the breakpoints across the alignment, we plotted a line graph of the relative frequency of the breakpoints. The K-means univariate-clustering algorithm as implemented in the ‘ckmeans.1d.dp’ R package was used to define hotspots for recombination³⁴. The gap statistic method implemented in the ‘factoextra’ R-package was employed to estimate the groups based on similarity in breakpoint positions obtained from the Simplot analysis³⁵. The recombination hotspot positions were then used to identify groups of three or more URFs with one or more similar breakpoint positions. Finally, to determine potential new CRFs, we performed a ML phylogenetic analysis using Garli (GTR nucleotide substitution model) to assess the epidemiological relatedness among the sequences³⁶.

Cluster analysis. A previously described BLAST approach was used to determine non-Nigerian subtype/CRF-specific reference sequences^{32,37}. Subtype/CRF-specific reference sequences and Nigerian sequences were aligned and ML phylogenies determined as described previously³². Nigerian transmission clusters were defined as clusters with aLRT-SH support ≥ 0.90 and composed of $\geq 80\%$ Nigerian sequences^{32,38}. Clusters of two sequences were defined as dyads, 3–14 sequences as networks, and >14 sequences as large clusters³⁹.

Dated phylogeographic analysis. The temporal signal in each dataset were assessed by estimating the regression between divergence and sampling dates in TempEst⁴⁰. For the main Nigerian transmission clusters with little or no temporal structure, we used substitution rate priors obtained by preliminary analysis of 150 randomly selected sequences sampled worldwide for each subtype/CRF. The evolutionary rates were estimated in BEAST v1.8.4⁴¹ using the SRD06 model⁴², a relaxed uncorrelated lognormal clock model^{43,44}, and the Bayesian Skygrid coalescent tree prior⁴⁵. For each transmission cluster, Markov chain Monte Carlo (MCMC) chains were run for 300 million steps, subsampling parameters and trees every 30000th step. BEAGLE library v.2 was used to improve computational time of likelihood calculations⁴⁶. Convergence was assessed using Tracer v.1.6. All parameters achieved convergence as determined by effective sample sizes (ESS) ≥ 100 ⁴⁷.

We used a Bayesian discrete phylogeographic approach with a MCMC length of 300 million steps in BEAST v1.8.4, sampling every 30000th step, to reconstruct the spatial dynamics of HIV-1 for the identified large clusters^{41,48}. Effective population sizes through time was inferred using the Bayesian Skygrid coalescent model⁴⁹. Growth rates were estimated using the parametric exponential growth rate coalescent model^{50,51}.

Symmetric and asymmetric continuous time Markov chain models were used to model the location exchange process⁵². The most parsimonious description of the location exchange rates was inferred using the Bayesian stochastic search variable selection (BSSVS) procedure⁵². A robust counting approach as implemented in BEAST was used to estimate the forward and reverse viral movement events between locations along the branches of time dated phylogenetic trees⁵³. Well-supported movements were summarized using SPREAD v1.0.7 based on a Bayes factors ≥ 3 ^{54–56}. The percentage of viral movements between locations was summarized using R⁵⁷. All files and scripts are available upon request.

Statistics. Linear by linear association test (LBL) was used to analyze trends over time, using IBM SPSS V22.0 Armonk, NY: IBM Corporation.

Ethics. All research was performed in accordance with relevant guidelines/regulations, and informed consent was obtained from all participants if appropriate. Specifically, the study utilized secondary data analysis from laboratory database of the National HIV program, which has approval from the Nigerian National Health Research Ethics Committee (NHREC Approval #NHREC/01/01/2007-14/08/2017).

Accession numbers for previously unpublished sequences used in this study. The DNA sequences of HIV-1 *pol* PR-RT regions determined as part of this study were submitted to GenBank under the following accession numbers: DQ990400 to DQ990455.

Results

CRF02_AG, CRF43_02G and subtype G were the major circulating strains in Nigeria. We analyzed 1442 HIV-1 *pol* sequences collected from four geopolitical zones in Nigeria between 1999 and 2013 (Table 1). The phylogenetic analysis showed that the CRF02_AG was the most common strain (44% of the analyzed sequences), followed by CRF43_02G (16%), subtype G (8%) and CRF06_cpx (4%). Three-hundred-and-twenty-eighty of the sequences (23%) were unique recombinant forms (URFs), whereas the remaining sequences were minor variants (each variant accounting for <2%).

Most sequences were from Abuja (697 sequences, 48%), followed by Lagos (346 sequences, 23%) and Jos (216 sequences, 15%). The distribution of different subtypes/CRFs varied within these regions/states with fewer CRF02_AG infections following a North East direction from Lagos (Fig. 1). Analysis of trends over time showed an overall decrease in the proportion of CRF02_AG infections in Nigeria (from 55% in 2006 to 38% in 2013, $p = 0.015$, LBL, Supplementary Fig. S1). Moreover, the analysis also showed an increase in the proportion of unique recombinant forms (URFs) from 16% to 32%, 2005–2013 ($p = 0.015$, LBL). The time trend analysis for other CRFs and or subtypes was non-significant.

Four potential recombination hotspots in the *pol* region. Of the 328 identified URFs, 209 constituted variations of the three dominating circulating strains in Nigeria (CRF02_AG, CRF43_02G, and subtype G; Supplementary Fig. 4). The remaining 119 URFs consisted of different combinations of various HIV-1 strains less common in Nigeria. We therefore performed a detailed recombination analysis on the 209 URFs consisting of CRF02_AG, CRF43_02G, and subtype G. One-hundred of the 209 URF sequences (48%) were from Abuja; 45 from Lagos (22%); 35 from Jos (17%); 10 from Maiduguri (5%); 9 from Kaduna (4%); seven from Ibadan (3%); and two from Adamawa (1%); and one from Enugu (<1%). These sequences were initially selected from the maximum likelihood phylogenetic trees if they branched off close to the root between two subtypes and had long branches. There were 655 breakpoint positions recorded among the 209 sequences, with some sequences having more than one breakpoint. These positions were plotted as a frequency plot of breakpoints along the alignment to identify hotspots for recombination (i.e. positions in the alignment where recombination breakpoints occur more frequently than the other positions, Supplementary Fig. S2). Alignment positions around 285–315 (HXB2 K03455 positions: 2538–2568), 503–534 (2756–2787), 720–775 (2973–3028) and 923–956 (3176–3209) were identified as potential recombination hotspots. To define these positions more precisely, we used the inter-quartile range (IQR) of the optimal univariate K-median clustering algorithm with the number of clusters determined by the gap-statistic method (Supplementary Fig. S3)⁵⁸. Recombination hotspot I: 294–312 (HXB2 K03455 positions: 2547–2565); II: 503–533 (2756–2786); III: 729–805 (2982–3058); and IV: 931–957 (3184–3210) (Supplementary Table S1). Of the 209 sequences, 139 (67%) had a recombination breakpoint in the hotspot I region; 104 (50%) in hotspot II region; 58 (28%) in hotspot III region; and 59 (28%) in hotspot IV region. The hotspot positions were unique independent recombination events from the original breakpoint positions as previously identified for the parental sequences of CRF43_02G (HXB2 K03455 positions: 1266, 3325, and 6097) and CRF02_AG (HXB2 K03455 positions: 2391, 3275, and 4175). We identified 37 URFs that could be divided into seven groups sharing

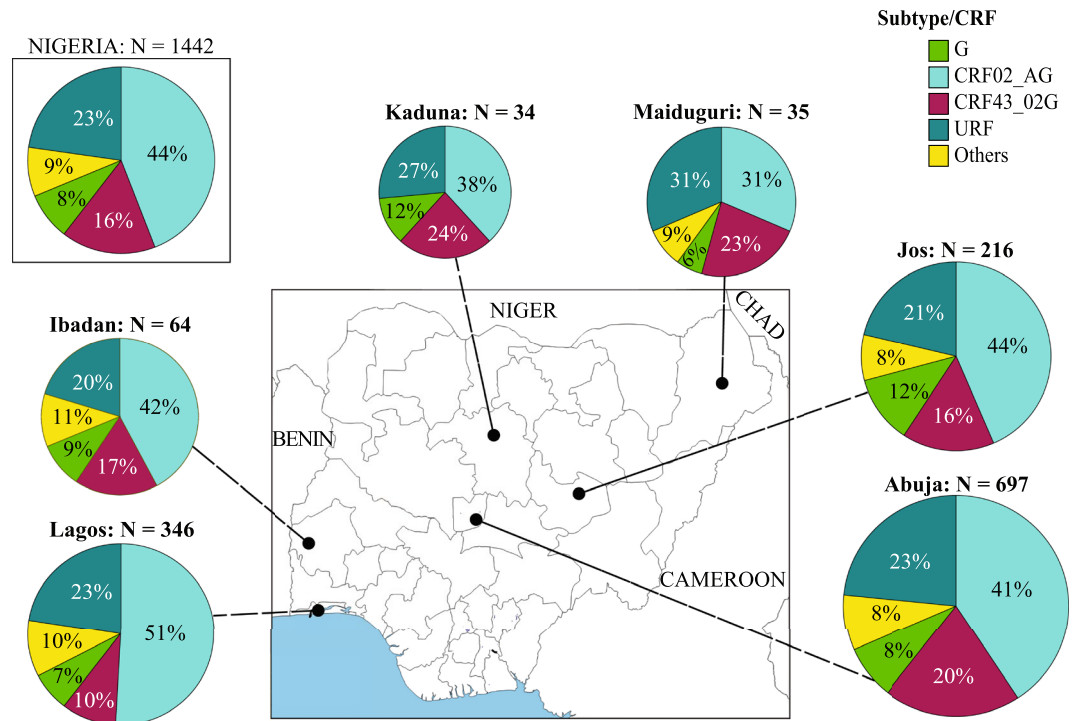


Figure 1. Prevalence of subtypes/CRFs among HIV-1 infected individuals collected from different locations in Nigeria. HIV-1 diversity in different zones/regions in Nigeria. The locations of the sampled regions as indicated on the map are connected to their respective diversity pie charts by a dot-headed dashed black line.

Subtype/CRF	Dyads ^a	Networks ^b	Large clusters ^c	Total
G	8 (4)	8 (2)	94 (1)	104
CRF02_AG	24 (12)	76 (12)	336 (6)	439
CRF43_02G	0 (0)	0 (0)	295 (1)	295
Total	32	84	725	841

Table 2. Number of clusters in the different subtype/CRF groups. ^aDyads: Clusters of 2 sequences. ^bNetworks: Clusters of 3–14 sequences. ^cLarge clusters: Clusters with more than 14 sequences. The number of clusters observed for the different groups are indicated in brackets.

breakpoint positions (each group consisting of 3–10 sequences, respectively, and thereby fulfilling the requirements as potentially novel CRFs; Supplementary Table S2). The majority of these sequences had been collected in Abuja (23/37, 62%).

Cluster analysis. To determine transmission clusters of the major circulating strains within Nigeria, we analyzed the three dominating forms: subtype G, CRF02_AG, and CRF43_02G. Analysis of 206 subtype G sequences (119 Nigerian and 87 non-Nigerian) showed four dyads, one network, and one large Nigerian subtype G cluster (consisting of 81 Nigerian and 11 non-Nigerian sequences, Table 2 and Supplementary Fig. S5). The 1161 CRF02_AG sequences (636 Nigerian and 555 non-Nigerian) formed 12 dyads, 12 networks and six clusters (Supplementary Fig. S6). Finally, analysis of the 295 CRF43_02G sequences (236 Nigerian and 59 non-Nigerian) showed that all the Nigerian sequences clustered monophyletically (SH-aLRT = 0.99, Supplementary Fig. 7), suggesting that CRF43_02G predominately circulates in Nigeria.

Dates of origin and evolutionary rates of the main Nigerian transmission clusters. To further dissect the Nigerian HIV-1 epidemic, we performed a detailed analysis of the five largest Nigerian clusters (one subtype G, three CRF02_AG, and one CRF43_02G). The time to most recent common ancestor (tMRCA) of the Nigerian subtype G cluster was 1975 (95%HPD: 1969–1982). The corresponding estimates for the CRF02_AG clusters were 1963 (95%HPD: 1948–1974), 1970 (95%HPD: 1960–1980) and 1960 (95%HPD: 1947–1974) for cluster 1, 2 and 3, respectively; and 1971 (95%HPD: 1952–1983) for the CRF43_02G cluster (Supplementary Fig. 8). The HIV-1 evolutionary rate for the Nigerian subtype G cluster was 2.1×10^{-3} substitutions/site/year (s/s/y, 95%HPD: $1.67\text{--}2.53 \times 10^{-3}$ s/s/y). The rates for the CRF02_AG clusters 1, 2 and 3 was slightly lower: 1.42×10^{-3} s/s/y (95% HPD: $1.07\text{--}1.85 \times 10^{-3}$ s/s/y) for cluster 1; 1.34×10^{-3} s/s/y (95%HPD: $1.02\text{--}1.72 \times 10^{-3}$ s/s/y) for

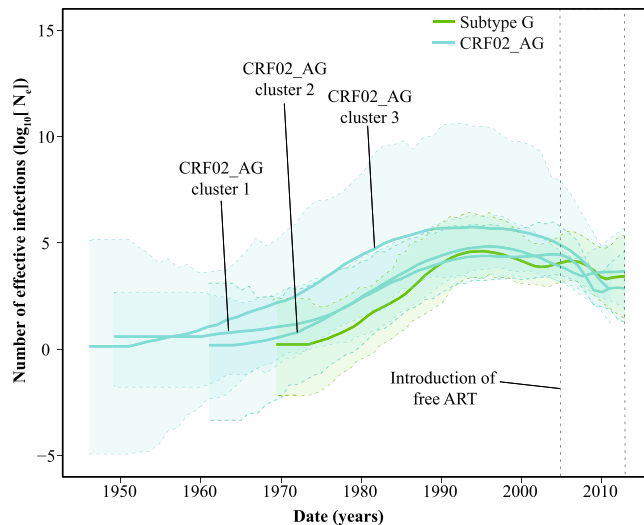


Figure 2. Skygrid plots for the different clusters. Phylodynamic analyses of HIV-1 *pol* gene for subtype G and CRF02_AG isolated in Nigeria. Bayesian Skygrid plots representing the changes in the effective population size of the virus (N_t) over time. The solid lines represent the estimated median log effective population size and the colour-coded shaded areas represent the 95%HPD intervals for the different clusters circulating in Nigeria. The gray-shaded area starts at the period when ART was introduced in Nigeria.

cluster 2; and 1.22×10^{-3} s/s/y (95%HPD: $0.91-1.57 \times 10^{-3}$ s/s/y) for cluster 3. Finally, the rate for the CRF43_02G cluster was estimated to 2.72×10^{-3} s/s/y (95%HPD: $1.59-3.73 \times 10^{-3}$ s/s/y) (Supplementary Fig. 8).

To assess the impact of sampling bias in our phylogeographic inference, we conducted a control analysis of the effects of potential over-representations of samples from particular regions in Nigeria (Table 1). Sequences were randomly selected based on population growth over time and HIV-1 prevalence in the different geographic regions. Information on the Nigerian population growth over time was obtained from the National Population Commission of Nigeria/National Bureau of Statistics for the population census, and information on HIV-1 prevalence was obtained from National Agency for the Control of AIDS¹⁶. In these analyses, the median tMRCAs of the Nigerian subtype G cluster was estimated to 1987 (95%HPD: 1982–1992); and the CRF02_AG clusters to 1974 (95%HPD: 1960–1983), 1972 (95%HPD: 1973–1981), 1961 (95%HPD: 1952–1979) for cluster 1, 2 and 3, respectively. Despite numerous attempts, the CRF43_02G control analysis did not reach sufficient temporal signal to converge.

Disentangling the demographic history of the main Nigerian transmission clusters. The analysis of the Nigerian subtype G epidemic showed that the number of effective infections, a proxy of HIV incidence, underwent a fast exponential growth between the 1970s and the mid-1990s, followed by a marginal decrease with minor fluctuations from the mid-1990s and onwards (Fig. 2)⁵⁹. Using the exponential growth model, the median growth rate was 0.3 per year (95%HPD: 0.18–0.42). The three clusters representing the CRF02_AG epidemic displayed a similar pattern with a slow increase in growth rate from the 1980s to 2000s. The median CRF02_AG growth rates were estimated to 0.22 per year (95%HPD: 0.13–0.32) for cluster 1; 0.18 per year (95%HPD: 0.10–0.26) for cluster 2; and 0.24 (95%HPD: 0.11–0.38) for cluster 3 (Fig. 2). Finally, the CRF43_02G cluster had a median growth rate of 0.39 per year (95%HPD: 0.23–0.55). The CRF43_02G cluster had low temporal signal that could not inform the complex non-parametric coalescent model and thus only an exponential model was used to capture the demographic dynamics (Fig. 3).

Phylogeographic dispersal of HIV-1 in the main Nigerian subepidemics. Next, we sought to investigate the spatiotemporal spread of HIV-1 in Nigeria using both asymmetric and symmetric discrete phylogeographic diffusion models. From the asymmetric analysis, various sample locations (cities) had well supported exchange rates that dominated the diffusion process of the different subepidemics. For subtype G, significant transition between Lagos and Ibadan (Bayes Factor [BF] = 241), Damaturu (BF = 109), Minna (BF = 174), and Jos (BF = 98) were identified. Frequently invoked rates were also identified between Abuja and Lagos (BF = 81), Kaduna (BF = 220), Ibadan (BF = 223), Damaturu (BF = 293) and Minna (BF = 138). Supported rates from Kaduna, Jos and Ibadan to the other locations within Nigeria were also found. Finally, the analysis also indicated significant international linkages to and from Lagos (BF = 74), Abuja (BF = 80), Kaduna (BF = 21), Jos (BF = 501), Damaturu (BF = 80), and Ibadan (BF = 207). The symmetric models gave similar results as the asymmetric model.

For the CRF02_AG cluster 1, Abuja was connected to Ibadan (BF = 105), Kaduna (BF = 687), and Jos (BF = 116); Lagos was connected to Ibadan (BF = 512), Kaduna (BF = 134) and Oyo (BF = 139); Kaduna was connected to Jos (BF = 207), Ibadan (BF = 371) and Oyo (BF > 1000). International linkage was also found for Lagos (BF = 639), Abuja (BF = 184), Ibadan (BF = 462) and Oyo (BF = 275). For CRF02_AG cluster 2, Lagos was connected to Jos (BF = 395), Maiduguri (BF = 231), Ibadan (BF > 1000) and Kaduna (BF = 383); Abuja to

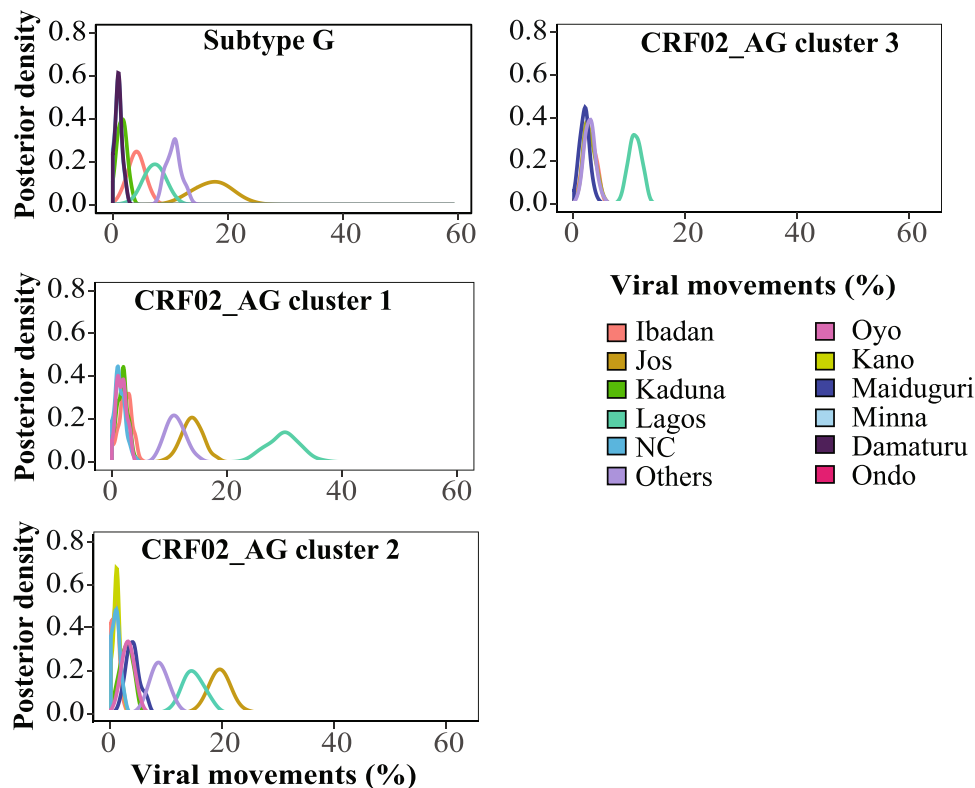


Figure 3. Estimated percentages of viral migration events from Abuja to each location in Nigeria. The density plot for the viral movements from the Abuja (most probable root location) to Lagos, Jos, Kaduna and other towns. Lagos had the highest percentage of viral movements from Abuja for the CRF02_AG cluster 1, and CRF02_AG cluster 3; whereas Jos had the highest percentage of viral movements from Abuja for CRF02_AG cluster 2 and subtype G.

Kaduna (BF = 155), Ibadan (BF = 110), Kano (BF = 543) and Oyo (BF = 58); and Ibadan to Kaduna, Jos, Oyo, Kano, and Maiduguri. For the CRF02_AG cluster 3, Lagos was connected to Abuja (BF = 118), Jos (BF = 87) and Maiduguri (BF = 56). Abuja was connected to Ibadan (BF = 45), Jos (BF = 44) and Maiduguri (BF = 29).

As described above, the CRF43_02G cluster had low temporal signal and could not inform the complex phylogenetic analysis. Finally, based on posterior probability support, the location of the spatial origin of subtype G, and CRF02_AG clusters was Abuja (posterior probability ≥ 0.98).

Rates of viral lineage migration. The rates of viral lineage migration within Nigeria were estimated using a robust counting approach. For subtype G, 42% (95%HPD: 35–48%) of all virus migrations originated in Abuja; of these 17% (95%HPD: 12–22%) were directed to Jos, 10% (95%HPD: 8–13%) to outside Nigeria, and 7% (95%HPD: 4–10%) to Lagos. Similar results were found for the CRF02_AG clusters, except for slightly higher migration rates from Abuja to Lagos for cluster 1 (30%, 95%HPD: 25–34%, Fig. 3). Overall, the CRF02_AG migration rates originating in Abuja was 61%, 52% and 21% for cluster 1, 2 and 3, respectively.

Discussion

In this study, we aimed to disentangle the history and spread of HIV-1 subtypes/CRFs in Nigeria using state-of-the-art phylogenetics to a large set of HIV-1 sequences collected in the five most populated Nigerian geopolitical zones. To the best of our knowledge, only one previous nationwide molecular epidemiology study from Nigeria have been presented⁶⁰. In this study, 55 HIV-1 *gp41* sequences (approximately 400 bp) collected throughout Nigeria during 1999 were analysed by neighbour-joining phylogenetics. However, the analysis was unable to distinguish between subtype A and CRF02_AG sequences – likely due to insufficient phylogenetic information in this relatively short gene fragment, as suggested by the authors. In-depth phylogenetic and phylodynamic analysis was therefore not possible in this study. In comparison, we analysed 1442 HIV-1 *pol* sequences (approximately 1000 bp long) collected throughout Nigeria between 1999 and 2013 using maximum-likelihood and Bayesian phylogenetics. Importantly, the HIV-1 *pol* gene (which is the most targeted genetic region for HIV-1 sequencing and by far most commonly analysed region in studies of HIV-1 molecular epidemiology) holds sufficient intrinsic genetic variability to permit the reconstruction of transmission histories by phylogenetic approaches^{26,32,61}. Hence, the larger size of our dataset, the relatively long time-window of sampling, and the use of state-of-the-art phylodynamic methods enabled us to perform the first in-depth nationwide HIV-1 phylodynamic study in Nigeria. In line with other previous studies from different geographic regions in Nigeria, we found that CRF02_AG, CRF43_02G and subtype G were the most prevalent strains (previous estimates ranging

from 19–60% for CRF02_AG and 22–50% for subtype G^{18–25,62}). The prevalence of CRF43_02G has only been reported in one previous study from Abuja (estimated to 19%)²¹. Non-representative sampling or local fluctuations between geographic regions in Nigeria could explain the large variations and discrepancies between studies. It could also be explained, in part, by the use of different subtyping tools between studies (which can differ in accuracy in assigning the correct subtype/CRF based on partial genome sequences)²⁷.

We found one large well-supported monophyletic cluster for both subtype G and CRF43_02G, respectively. Each of these clusters harbored the majority of Nigerian sequences, suggesting single strain introductions that grew out to dominate the Nigerian HIV-1 epidemic. Previous studies have used different nomenclatures of subtype G. A comparison of accession numbers showed that our subtype G cluster is partly consistent with the subtype G' cluster identified by Chaplin *et al.*²². Further dissection of the subtype G' cluster showed that approximately half of the sequences belonged to subtype G, whereas the other half clustered with the CRF43_02G sequences in our analysis. In a more recent study, Delatorre *et al.* suggested a nomenclature based on geographic dissemination⁶³. The clustering pattern of the West African strain defined as subtype G_{WA-II} was largely overlapping with the Nigerian subtype G cluster in our analysis⁶³. We estimated the date of origin for the subtype G cluster to 1975 (95%HPD: 1969–1982), which is consistent with the origin of the subtype G_{WA-II} cluster that was estimated to 1979 [95%HPD: 1973–1984]⁶³. In addition, the estimated date of origin also fits with the fact that the first AIDS cases were identified in Nigeria in 1986, approximately one decade after the introduction of this virus strain¹⁶. The CRF43_02G was first described and isolated in Saudi Arabia in 2008⁶⁴. Interestingly, the G_{WA-I} cluster determined by Delatorre *et al.* consisted of 140/168 (83%) sequences from Nigeria⁶³. In our analysis, more than half of the G_{WA-I} sequences clustered with the CRF43_02G strain. Spatiotemporal analysis of the G_{WA-I} cluster indicated that this strain emerged in Nigeria in the mid-1970s⁶³. This is consistent with our estimate of 1971 (95%HPD: 1952–1983) for the CRF43_02G cluster. Moreover, the majority of CRF43_02G sequences in Genbank are of Nigerian origin (www.hiv.lanl.gov), and both the putative parental strains of CRF43_02G in Nigeria (CRF02_AG and subtype G) are highly prevalent. Altogether, this suggests that CRF43_02G originated in central West Africa, in or close to Nigeria.

In contrast to subtype G and CRF43_02G, where only one of the respective introductions resulted in major subepidemics, three large Nigerian subepidemics were found for CRF02_AG. The underlying mechanisms for this strain-specific difference are unknown. However, the CRF02_AG strain is the most prevalent strain in West Africa, and a previous study of the West Central African CRF02_AG epidemic estimated that this strain originated in the Democratic Republic of the Congo (DRC) in 1973 (95%HPD: 1972–1975)⁶⁵. This is in line with a previous study by Abecasis *et al.* that estimated the CRF02_AG date of origin to 1976 (95%HPD: 1971–1981)⁶⁶. However, a more detailed analysis of the CRF02_AG strain by Mir *et al.* indicated that seven subvariants of CRF02_AG are circulating (as determined by a combination of phylogenetics and geographic dissemination)⁶⁷. The seven subvariants of CRF02_AG were estimated to have originated between 1967 and 2003 (combined 95%HPD: 1961–2005). More specifically, the analysis indicated that the oldest strain was the West African CRF02_AG strain (introduced 1967 [95%HPD: 1961–1974]). The three CRF02_AG clusters identified in the current study were estimated to have originated between 1963 and 1970 (combined 95%HPD: 1948–1980). Although the 95%HPD intervals are overlapping with the estimates suggesting the origin to be the DRC, they are on the lower side. Moreover, the CRF02_AG strain was first isolated in Nigeria in 1996⁶⁸. Due to the somewhat conflicting data presented above, further analysis is needed to provide solid evidence of the geographical origin of the CRF02_AG strain.

Bayesian demographic analysis indicated an increase in the number of effective infections 1970–1995 in all the analysed clusters. Interestingly, a rapid population growth occurred in Nigeria during the same period (from 56 to 108 million people, www.worldometers.info/world-population/nigeria-population). This increase was followed by a decline in number of effective infections that coincided with the introduction of free ART in Nigeria in 2006, which resulted in a decrease in HIV-1 prevalence from 6% to 3% in Nigeria the following years^{2,16}. Moreover, our analysis suggested that urban areas like Abuja and Lagos were the major hubs of HIV-1 transmission in Nigeria. This is in line with previous reports from West Africa⁶⁹. However, it should be noted that the majority of the sequences were collected in Abuja, and potential transmissions from other cities may therefore not have been captured in our analysis.

The recombination analysis indicated several URFs and potentially novel CRFs. Recombination does not occur randomly on the HIV-1 genome as its frequency varies along its length with several so-called hotspots for recombination^{70,71}. One main hotspot for recombination (found in 105 sequences) was close to the *pol PR-RT* border (positions 2547–2565 in the HIV-1 HXB2 reference genome, K03455). This hotspot has previously been reported in a study on HIV-1 subtype B⁷². The hotspot positions II and IV have also been reported previously⁷³. To date, eighteen distinct subtype G-related CRFs have been described (www.hiv.lanl.gov). We identified seven groups of URFs with similar recombination breakpoint patterns. These could represent novel CRFs or second-generation recombinants circulating in Nigeria. However, full-length genome sequencing is needed to confirm whether these groups represent similar URFs or previously unknown subtype G-related CRFs.

In summary, this is the first in-depth HIV-1 phylodynamic study based on a nationwide set of HIV-1 sequences from Nigeria. We found a high number of URFs and potential new CRFs, and our analyses suggested that HIV-1 first emerged and expanded within large urban centers before migrating to smaller rural areas. The number of effective infections declined in the early 2000s, coinciding with the introduction of free ART in Nigeria. This study increases our understanding of the Nigerian HIV-1 epidemic and may inform HIV-1 intervention strategies to reduce the spread of HIV-1 in Nigeria.

Received: 5 August 2019; Accepted: 5 February 2020;

Published online: 26 February 2020

References

- Gottlieb, M. S. *et al.* Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. *The New England journal of medicine* **305**, 1425–1431 (1981).
- UNAIDS. UNAIDS, Global Reports - UNAIDS report on the global AIDS epidemic 2016. (2017).
- Palm, A. A. *et al.* Faster progression to AIDS and AIDS-related death among seroincident individuals infected with recombinant HIV-1 A3/CRF02_AG compared with sub-subtype A3. *The Journal of infectious diseases* **209**, 721–728 (2014).
- Naidoo, V. L. *et al.* Mother-to-Child HIV Transmission Bottleneck Selects for Consensus Virus with Lower Gag-Protease-Driven Replication Capacity. *Journal of virology* **91**(2017).
- Laheer, F. *et al.* HIV Controllers Exhibit Enhanced Frequencies of Major Histocompatibility Complex Class II Tetramer(+) Gag-Specific CD4(+) T Cells in Chronic Clade C HIV-1 Infection. *Journal of virology* **91** (2017).
- Kiwanuka, N. *et al.* HIV-1 viral subtype differences in the rate of CD4+ T-cell decline among HIV seroincident antiretroviral naive persons in Rakai district, Uganda. *J Acquir Immune Defic Syndr* **54**, 180–184 (2010).
- Baeten, J. M. *et al.* HIV-1 subtype D infection is associated with faster disease progression than subtype A in spite of similar plasma HIV-1 loads. *The Journal of infectious diseases* **195**, 1177–1180 (2007).
- Senkaali, D. *et al.* The relationship between HIV type 1 disease progression and V3 serotype in a rural Ugandan cohort. *AIDS Res Hum Retroviruses* **20**, 932–937 (2004).
- Kaleebu, P. *et al.* Relationship between HIV-1 Env subtypes A and D and disease progression in a rural Ugandan cohort. *AIDS* **15**, 293–299 (2001).
- Kanki, P. J. *et al.* Human immunodeficiency virus type 1 subtypes differ in disease progression. *The Journal of infectious diseases* **179**, 68–73 (1999).
- Esbjornsson, J. *et al.* Frequent CXCR4 tropism of HIV-1 subtype A and CRF02_AG during late-stage disease—indication of an evolving epidemic in West Africa. *Retrovirology* **7**, 23 (2010).
- Mild, M. *et al.* High inpatient HIV-1 evolutionary rate is associated with CCR5-to-CXCR4 coreceptor switch. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases* **19**, 369–377 (2013).
- Laga, M. Effect of HIV prevention in key populations: evidence accumulates, time to implement. *Lancet Glob Health* **1**, e243–244 (2013).
- Granich, R. *et al.* Trends in AIDS Deaths, New Infections and ART Coverage in the Top 30 Countries with the Highest AIDS Mortality Burden, 1990–2013. *PLoS one* **10**, e0131353 (2015).
- UNAIDS. UNAIDS DATA 2018. (2018).
- (NACA), N.A.F.T.C.O.A. Global AIDS Response Country Progress report Nigeria GARPR 2015. (2015).
- Ajoge, H. O. *et al.* Drug resistance pattern of HIV type 1 isolates sampled in 2007 from therapy-naive pregnant women in North-Central Nigeria. *AIDS Res Hum Retroviruses* **28**, 115–118 (2012).
- Imade, G. E. *et al.* Short communication: Transmitted HIV drug resistance in antiretroviral-naive pregnant women in north central Nigeria. *AIDS Res Hum Retroviruses* **30**, 127–133 (2014).
- Volz, E. M. *et al.* Phylodynamic analysis to inform prevention efforts in mixed HIV epidemics. *Virus evolution* **3**, vex014 (2017).
- Charurat, M. *et al.* Characterization of acute HIV-1 infection in high-risk Nigerian populations. *The Journal of infectious diseases* **205**, 1239–1247 (2012).
- Diallo, K., Zheng, D. P., Rottinghaus, E. K., Bassey, O. & Yang, C. Viral Genetic Diversity and Polymorphisms in a Cohort of HIV-1-Infected Patients Eligible for Initiation of Antiretroviral Therapy in Abuja, Nigeria. *AIDS Res Hum Retroviruses* **31**, 564–575 (2015).
- Chaplin, B. *et al.* Impact of HIV type 1 subtype on drug resistance mutations in Nigerian patients failing first-line therapy. *AIDS Res Hum Retroviruses* **27**, 71–80 (2011).
- Ojesina, A. I. *et al.* Subtype-specific patterns in HIV Type 1 reverse transcriptase and protease in Oyo State, Nigeria: implications for drug resistance and host response. *AIDS Res Hum Retroviruses* **22**, 770–779 (2006).
- Ajoge, H. O. *et al.* Genetic characteristics, coreceptor usage potential and evolution of Nigerian HIV-1 subtype G and CRF02_AG isolates. *PLoS one* **6**, e17865 (2011).
- Hamers, R. L. *et al.* HIV-1 drug resistance in antiretroviral-naive individuals in sub-Saharan Africa after rollout of antiretroviral therapy: a multicentre observational study. *The Lancet. Infectious diseases* **11**, 750–759 (2011).
- Hassan, A. S., Pybus, O. G., Sanders, E. J., Albert, J. & Esbjornsson, J. Defining HIV-1 transmission clusters based on sequence data. *Aids* **31**, 1211–1222 (2017).
- Pineda-Pena, A. C. *et al.* Automated subtyping of HIV-1 genetic sequences for clinical and surveillance purposes: performance evaluation of the new REGA version 3 and seven other tools. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases* **19**, 337–348 (2013).
- Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)* **23**, 2947–2948 (2007).
- Tamura, K., Stecher, G., Peterson, D., Filipiński, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular biology and evolution* **30**, 2725–2729 (2013).
- Zwickl, D. J. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. *PhD thesis* (2006).
- Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic biology* **59**, 307–321 (2010).
- Esbjornsson, J. *et al.* HIV-1 transmission between MSM and heterosexuals, and increasing proportions of circulating recombinant forms in the Nordic Countries. *Virus evolution* **2**, vew010 (2016).
- Lole, K. S. *et al.* Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *Journal of virology* **73**, 152–160 (1999).
- Wang, H. & Song, M. Ckmeans.1d.dp: Optimal k-means Clustering in One Dimension by Dynamic Programming. *The R journal* **3**, 29–33 (2011).
- Kassambara, A. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. (2017).
- Robertson, D. L. *et al.* HIV-1 Nomenclature Proposal. *Science* **288**, 55–55 (2000).
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Journal of molecular biology* **215**, 403–410 (1990).
- Anisimova, M., Gil, M., Dufayard, J. F., Dessimoz, C. & Gascuel, O. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Systematic biology* **60**, 685–699 (2011).
- Aldous, J. L. *et al.* Characterizing HIV transmission networks across the United States. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America* **55**, 1135–1143 (2012).
- Rambaut, A., Lam, T. T., Max Carvalho, L. & Pybus, O. G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus evolution* **2**, vew007–vew007 (2016).
- Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular biology and evolution* **29**, 1969–1973 (2012).
- Shapiro, B., Rambaut, A. & Drummond, A. J. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Molecular biology and evolution* **23**, 7–9 (2006).
- Kishino, H., Thorne, J. L. & Bruno, W. J. Performance of a divergence time estimation method under a probabilistic model of rate evolution. *Molecular biology and evolution* **18**, 352–361 (2001).

44. Thorne, J. L., Kishino, H. & Painter, I. S. Estimating the rate of evolution of the rate of molecular evolution. *Molecular biology and evolution* **15**, 1647–1657 (1998).
45. Drummond, A. J., Rambaut, A., Shapiro, B. & Pybus, O. G. Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular biology and evolution* **22**, 1185–1192 (2005).
46. Ayres, D. L. *et al.* BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. *Systematic biology* **61**, 170–173 (2012).
47. Rambaut, A., Suchard, M. A., Xie, D. & Drummond, A. J. Tracer v1.6, Available from <http://tree.bio.ed.ac.uk/software/tracer/>. (2014).
48. Lemey, P., Rambaut, A., Welch, J. J. & Suchard, M. A. Phylogeography takes a relaxed random walk in continuous space and time. *Molecular biology and evolution* **27**, 1877–1885 (2010).
49. Baele, G. *et al.* Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular biology and evolution* **29**, 2157–2167 (2012).
50. Rogers, A. R. & Harpending, H. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular biology and evolution* **9**, 552–569 (1992).
51. Slatkin, M. & Hudson, R. R. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **129**, 555–562 (1991).
52. Lemey, P., Rambaut, A., Drummond, A. J. & Suchard, M. A. Bayesian phylogeography finds its roots. *PLoS computational biology* **5**, e1000520 (2009).
53. Minin, V. N. & Suchard, M. A. Counting labeled transitions in continuous-time Markov models of evolution. *Journal of mathematical biology* **56**, 391–412 (2008).
54. Bielejec, F., Rambaut, A., Suchard, M. A. & Lemey, P. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics (Oxford, England)* **27**, 2910–2912 (2011).
55. Drummond, A. J. & Bouckaert, R. R. Bayesian evolutionary analysis with BEAST. (Cambridge University Press., 2017).
56. Kass, R. E. & Raftery, A. E. Bayes Factors. *Journal of the American Statistical Association* **90**, 773–795 (1995).
57. Team, R.C. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>. (2013).
58. Robert, T., Guenther, W. & Trevor, H. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **63**, 411–423 (2001).
59. Gill, M. S. *et al.* Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. *Molecular biology and evolution* **30**, 713–724 (2013).
60. Agwale, S. M. *et al.* Molecular surveillance of HIV-1 field strains in Nigeria in preparation for vaccine trials. *Vaccine* **20**, 2131–2139 (2002).
61. Hue, S., Clewley, J. P., Cane, P. A. & Pillay, D. HIV-1 pol gene variation is sufficient for reconstruction of transmissions in the era of antiretroviral therapy. *AIDS* **18**, 719–728 (2004).
62. Peeters, M. *et al.* Predominance of subtype A and G HIV type 1 in Nigeria, with geographical differences in their distribution. *AIDS Res Hum Retroviruses* **16**, 315–325 (2000).
63. Delatorre, E., Mir, D. & Bello, G. Spatiotemporal dynamics of the HIV-1 subtype G epidemic in West and Central Africa. *PLoS one* **9**, e98908 (2014).
64. Yamaguchi, J. *et al.* Identification of new CRF43_02G and CRF25_cpx in Saudi Arabia based on full genome sequence analysis of six HIV type 1 isolates. *AIDS Res Hum Retroviruses* **24**, 1327–1335 (2008).
65. Faria, N. R. *et al.* Phylodynamics of the HIV-1 CRF02_AG clade in Cameroon. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases* **12**, 453–460 (2012).
66. Abecasis, A. B., Vandamme, A. M. & Lemey, P. Quantifying differences in the tempo of human immunodeficiency virus type 1 subtype evolution. *Journal of virology* **83**, 12917–12924 (2009).
67. Mir, D. *et al.* Phylodynamics of the major HIV-1 CRF02_AG African lineages and its global dissemination. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases* **46**, 190–199 (2016).
68. Howard, T. M. & Rasheed, S. Genomic structure and nucleotide sequence analysis of a new HIV type 1 subtype A strain from Nigeria. *AIDS Res Hum Retroviruses* **12**, 1413–1425 (1996).
69. Esbjornsson, J., Mild, M., Mansson, F., Norrgren, H. & Medstrand, P. HIV-1 molecular epidemiology in Guinea-Bissau, West Africa: origin, demography and migrations. *PLoS one* **6**, e17025 (2011).
70. Zhuang, J. *et al.* Human immunodeficiency virus type 1 recombination: rate, fidelity, and putative hot spots. *Journal of virology* **76**, 11273–11282 (2002).
71. Magiorkinis, G. *et al.* In vivo characteristics of human immunodeficiency virus type 1 intersubtype recombination: determination of hot spots and correlation with sequence similarity. *J. Gen Virol.* **84**, 2715–2722 (2003).
72. Galli, A. *et al.* Recombination analysis and structure prediction show correlation between breakpoint clusters and RNA hairpins in the pol gene of human immunodeficiency virus type 1 unique recombinant forms. *J. Gen Virol.* **89**, 3119–3125 (2008).
73. Smyth, R. P. *et al.* Identifying recombination hot spots in the HIV-1 genome. *Journal of virology* **88**, 2891–2902 (2014).

Acknowledgements

We thank all the study participants, the collaborating health centers in Nigeria, Federal Ministry of Health and the National Agency for the Control of AIDS (NACA) in Nigeria. The study was supported by the Swedish Research Council (No. 350-2012-6628 and 2016-01417), Swedish Society of Medical Research (SA-2016), the Medical Faculty at Lund University, the US NIH R01 AI147331-01 (RAI147331A), and the US Centers for Disease Control and Prevention to the Institute of Human Virology Nigeria (5U2GGH000925). Open access funding provided by Lund University.

Author contributions

J.N., N.N. and J.E. interpreted the data and were responsible for the overall study design. N.N. and J.E. were responsible for the overall project coordination. H.R., B.C., N.N., P.K., P.D., A.A., M.C. were medically and organizationally responsible for the clinical sites and collected key epidemiological data of the study participants. J.N., N.R.F., N.N. and J.E. analyzed the data and contributed in statistical analyses. J.N. and J.E. wrote the manuscript. All authors read and approved the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-59944-x>.

Correspondence and requests for materials should be addressed to J.E.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020